

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-259036

(43)Date of publication of application : 03.10.1997

(51)Int.Cl.

G06F 12/08

G06F 12/08

(21)Application number : 09-040832

(71)Applicant : INTERNATL BUSINESS MACH CORP <IBM>

(22)Date of filing : 25.02.1997

(72)Inventor : RAJESSHU VIKUFUBAI PATERU
SOMIA MARRIC

(30)Priority

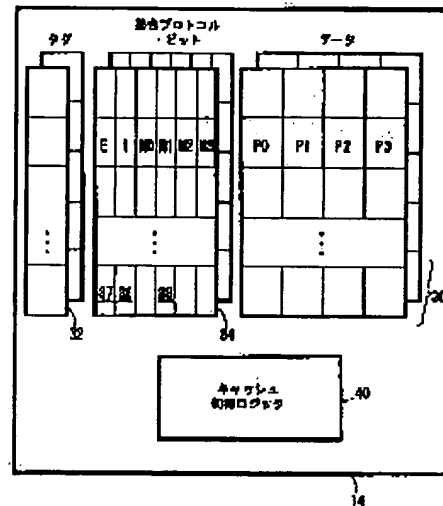
Priority number : 96 616612 Priority date : 15.03.1996 Priority country : US

(54) WRITE-BACK CACHE AND METHOD FOR MAINTAINING CONSISTENCY IN WRITE-BACK CACHE

(57)Abstract:

PROBLEM TO BE SOLVED: To maintain the consistency inside a write-back cache memory by providing plural cache lines, a matching indicator and cache control logic and making the cache control logic write only an identified byte set to a related memory at the time of replacing the cache lines inside the write-back cache memory.

SOLUTION: The plural cache lines 30 including the plural byte sets store data related to addresses inside the related memory. Then, the matching indicator identifies the byte set provided with the data different from the data stored in the corresponding address inside the related memory among the plural byte sets inside the cache lines 30. The cache control logic 40 writes only the identified byte set to the related memory at the time of replacing the selected way of the cache lines 30.



LEGAL STATUS

[Date of request for examination] 11.12.1998

[Date of sending the examiner's decision of rejection] 06.11.2001

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-259036

(43) 公開日 平成9年(1997)10月3日

(51) Int.Cl. ⁴	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 12/08		7623-5B	G 0 6 F 12/08	B
	3 1 0	7623-5B		3 1 0 Z

審査請求 未請求 請求項の数17 O L (全 9 頁)

(21) 出願番号 特願平9-40832

(22) 出願日 平成9年(1997)2月25日

(31) 優先権主張番号 08/616612

(32) 優先日 1996年3月15日

(33) 優先権主張国 米国 (U S)

(71) 出願人 390009531

インターナショナル・ビジネス・マシー
ズ・コーポレーション

INTERNATIONAL BUSIN
ESS MASCHINES CORPO
RATION

アメリカ合衆国10504、ニューヨーク州
アーモンク (番地なし)

(72) 発明者 ラジェッシュ・ヴィクフパイ・パテル

アメリカ合衆国78748、テキサス州オー
スティン シルク オーク コープ 9313

(74) 代理人 弁理士 合田 潔 (外2名)

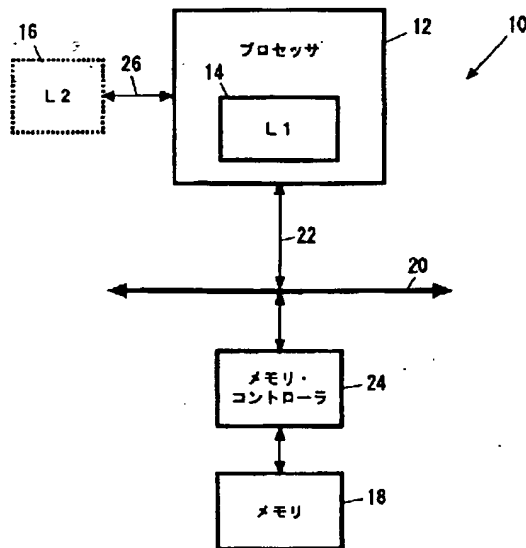
最終頁に続く

(54) 【発明の名称】 ライトバックキャッシュおよびライトバックキャッシュ内で整合性を維持する方法

(57) 【要約】

【課題】 ライトバックキャッシュメモリと、そのメモリ内で整合性を維持する方法を提供すること。

【解決手段】 ライトバックキャッシュメモリは、関連づけられるメモリ内のアドレスに関連したデータを記憶する複数のキャッシュラインを含む。キャッシュラインの各々は複数のバイト集合を含む。さらに、ライトバックキャッシュメモリは整合標識を含み、この整合標識は、キャッシュラインに含まれる複数のバイト集合の中のどのバイト集合が、関連づけられるメモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むかを識別する。さらに、ライトバックキャッシュメモリはキャッシュ制御ロジックを含み、このロジックは、ライトバックキャッシュメモリ内の特定のキャッシュラインが置換されるとき、識別されたバイト集合のみを、関連づけられるメモリへ書き込むので、メモリアクセスが少なくなる。



【特許請求の範囲】

【請求項1】 複数のバイト集合を含む複数のキャッシュラインであって、関連づけられるメモリ内のアドレスに関連づけられたデータを記憶するものと、上記キャッシュライン内の上記複数のバイト集合の中で、上記関連づけられるメモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むバイト集合を識別する整合標識と、上記キャッシュラインの特定のものを置換するとき、上記識別されたバイト集合のみを上記関連づけられるメモリへ書き込むキャッシュ制御ロジックと、を具備するライトバックキャッシュメモリ。

【請求項2】 請求項1において、上記整合標識が複数のビットを含み、該複数のビットの各々が上記複数のバイト集合の1つと関連づけられているライトバックキャッシュメモリ。

【請求項3】 請求項1において、上記バイト集合の各々が、ライトバックキャッシュメモリによって、上記関連づけられるメモリへ同時に書き込むことができる最大数のバイトを含むことを特徴とするライトバックキャッシュメモリ。

【請求項4】 請求項1において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有し、該複数のウエイの各々が、その上記複数のバイト集合の中で変更されたバイト集合を識別する整合標識を有することを特徴とするライトバックキャッシュメモリ。

【請求項5】 請求項1において、プロセッサと関連づけられた上位レベルキャッシュが含まれていることを特徴とするライトバックキャッシュメモリ。

【請求項6】 請求項1において、上記キャッシュ制御ロジックが、LRUアルゴリズムに従って上記キャッシュラインの特定のものを置換することを特徴とするキャッシュメモリ。

【請求項7】 プロセッサと、上記プロセッサへ接続されたメモリと、上記メモリへ接続されたライトバックキャッシュメモリとを具備し、該ライトバックキャッシュメモリには、複数のバイト集合を含む複数のキャッシュラインであって、上記メモリ内のアドレスと関連づけられたデータを記憶するものと、上記キャッシュライン内の上記複数のバイト集合の中で、上記メモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むバイト集合を識別する整合標識と、上記ライトバックキャッシュメモリ内の特定のキャッシュラインを置換するとき、上記識別されたバイト集合のみを上記メモリへ書き込むキャッシュ制御ロジックと、が含まれていることを特徴とするデータ処理システム。

【請求項8】 請求項7において、上記整合標識が複数のビットを含み、該複数のビットの各々が上記複数のバイト集合の1つと関連づけられていることを特徴とするデータ処理システム。

【請求項9】 請求項7において、上記バイト集合の各々が、上記ライトバックキャッシュメモリによって上記メモリへ同時に書き込むことができる最大数のバイトを含むことを特徴とするデータ処理システム。

【請求項10】 請求項7において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有し、該複数のウエイの各々が、そのバイト集合の中の変更されたものを識別する整合標識を有することを特徴とするデータ処理システム。

【請求項11】 請求項7において、上記ライトバックキャッシュメモリが上位レベルのキャッシュを含むことを特徴とするデータ処理システム。

【請求項12】 請求項7において、上記キャッシュ制御ロジックがLRUアルゴリズムに従って上記特定のキャッシュラインを置換することを特徴とするデータ処理システム。

【請求項13】 複数のキャッシュラインを有するライトバックキャッシュメモリの中で整合性を維持する方法であって、関連づけられるメモリ内のアドレスに関連づけられたデータを、上記複数のキャッシュラインの中の特定のキャッシュラインに記憶するステップと、上記特定のキャッシュライン内の複数のバイト集合の中で、上記関連づけられるメモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むバイト集合を識別するステップと、上記ライトバックキャッシュメモリ内の上記特定のキャッシュラインを置換するとき、上記識別されたバイト集合のみを上記関連づけられるメモリへ書き込むステップとより成る、ライトバックキャッシュメモリ内で整合性を維持する方法。

【請求項14】 請求項13において、上記ライトバックキャッシュメモリがそれに関連づけられた複数のビットを有し、上記識別ステップでは、上記複数のビットの中で、識別されるバイト集合に関連づけられたビットがセットされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

【請求項15】 請求項13において、上記ライトバックキャッシュメモリと上記関連づけられるメモリがバスによって接続され、上記書き込みステップでは、識別されたバイト集合のためのバストランザクションのみが必要とされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

【請求項16】 請求項13において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有し、該複数のウエイの各々が、その上記複数のバイト集合の中の変更されたものを識別する整合標識を有し、上記識別ステップでは、上記特定のキャッシュラインの選択されたウエイに含まれる上記複数のバイト集合の中の識別されるべきバイト集合に関連したビットがセットされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

【請求項17】請求項13において、上記特定のキャッシュラインがLRUアルゴリズムに従って置換されることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は一般的にはキャッシュメモリに関し、具体的にはライトバック(write-back)キャッシュメモリ内で整合性を維持することに関する。さらに具体的には、本発明は、キャッシュライン内に含まれるデータが細分されたラインサイズで更新されるライトバックキャッシュメモリと、このライトバックキャッシュメモリ内で整合性を維持する方法に関する。

【0002】

【従来の技術】キャッシュは高価な小型の高速メモリであり、関連づけられるメモリ内に記憶されたデータへのアクセス時間を改善し、プロセッサの遅延を減少させるために、データ処理システムで普通に使用されている。典型的には、キャッシュは多数のキャッシュラインから構成され、各キャッシュラインはいくつかのデータバイトを含む。メモリ内に記憶されたデータは、そのデータに関連づけられたメモリアドレスのインデックス部分を使用してキャッシュへマップされ、同じインデックス部分をもつ複数のメモリアドレスが同じキャッシュラインにマップされるようになっている。特定のメモリアドレスに関連づけられたキャッシュ内のデータは、アドレスタグによって、同じインデックス部分をもつ他のアドレスに関連づけられたデータから識別される。アドレスタグは、典型的には高順位のアドレスビットであり、キャッシュ内のデータと関連づけて記憶されている。同じインデックス部分をもつアドレスに関連づけられたデータ相互間の衝突を最小にするため、多くのキャッシュは各キャッシュラインについて複数のウェイすなわち記憶位置を備えている。

【0003】プロセッサから要求されたデータがキャッシュ内に存在しないとき、キャッシュミスが起こり、要求されたデータはメモリからフェッチされる。要求されたデータがキャッシュ内にあるようにするため、要求されたデータがマップされたキャッシュラインのウェイに存在するデータは、しばしば置換すなわちカーストアウト(cast-out)される必要がある。キャッシュが複数のウェイを有するとき、置換されるキャッシュ位置は、プロセッサによって高い確率で要求されるデータがカーストアウトされないように、典型的には、LRU(least recently used)アルゴリズムのようなアルゴリズムを使用して選択される。同様に、キャッシュ位置を検査して、キャッシュ位置に含まれるデータが、対応するメモリアドレス内に記憶された最新のデータと整合性をもたない場合、キャッシュ位置

を置換すなわちプッシュアウト(push-out)することができる。

【0004】キャッシュ内のデータとメモリ内のデータとの整合性は、キャッシュ整合プロトコルによって実効化される。このプロトコルは典型的にはプロセッサまたはキャッシュコントローラによって実施される。データアクセスがライトスルー(write-through)として指定されると、キャッシュ内データの更新は自動的にメモリへ書かれ、したがってデータアクセスの完了時にはキャッシュとメモリの整合性が維持される。他方、ライトバック(write-back)キャッシュでは、キャッシュラインの置換時またはバス活動が少ないときにコピーバック動作を実行することによってのみ、データの整合性が維持される。ライトバックキャッシュでは、バスの使用が少なくなるので、パフォーマンスの点からいえば、ソフトウェアとハードウェア環境によって低い整合性が許される場合には、一般的にライトバックキャッシュの方がライトスルーキャッシュよりも優っている。

【0005】キャッシュの整合および置換動作は、典型的にはライン単位で実行される。したがって、通常のライトバックキャッシュでは、各キャッシュラインの各ウェイに1つのビットが関連づけられており、このビットは、データバイトがキャッシュへ書き込まれた後でウェイ内のデータバイトが変更されたかどうかを示している。変更されたデータを含むウェイが置換されるとき、コピーバック動作が呼び出され、置換されるウェイの各バイトを、対応するメモリアドレスへ書き込む。典型的には、キャッシュラインの長さは、キャッシュが接続されているバスのバンドワイズ(bandwidth)よりも数倍長いので、通常、コピーバック動作が完了するまでに、多くのマルチサイクルバストランザクションすなわちビート(beat)を必要とする。

【0006】ライトバックキャッシュは、ライトスルーアーキテクチャのキャッシュと比較して優れたパフォーマンスを示すが、通常のライトバックキャッシュは、変更されないデータバイトも不必要にメモリへ書かれるために、置換されるウェイ内にあるデータバイトの一部だけが変更されるとき、バスのバンドワイズを不必要に消費する。変更されないデータバイトのライトバックは、バスのバンドワイズを消費してシステムの全体的パフォーマンスを低下させるだけでなく、特にバッテリー電源のコンピューティング環境(たとえば、携帯用コンピュータ)で問題となるが、キャッシュの能動化とバスの駆動時に電力を浪費する。したがって、変更されていないデータはメモリへ書き戻さないようなライトバックキャッシュが望まれる。

【発明が解決しようとする課題】

【0007】本発明の目的は、キャッシュメモリを提供することである。

【0008】本発明の他の目的は、ライトバックキャッシュメモリ内で整合性を維持するための改善された方法を提供することである。

【0009】さらに、本発明の他の目的は、各キャッシュライン内に含まれるデータの変更を、細分されたラインサイズで行うライトバックキャッシュメモリと、そのライトバックキャッシュメモリ内で整合性を維持する方法を提供することである。

【課題を解決するための手段】

【0010】前記の目的を達成するライトバックキャッシュメモリと、そのライトバックキャッシュメモリ内で整合性を維持する方法の手段は次のとおりである。すなわち、ライトバックキャッシュメモリは、関連づけられるメモリ内のアドレスに関連したデータを記憶する多数のキャッシュラインを含む。キャッシュラインの各々は複数のバイト集合を含む。さらに、ライトバックキャッシュメモリは整合標識を含み、この整合標識は、キャッシュライン内の複数のバイト集合の中でどのバイト集合が、関連づけられるメモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むかを識別する。さらに、ライトバックキャッシュメモリはキャッシュ制御ロジックを含む。このキャッシュ制御ロジックは、ライトバックキャッシュメモリ内の特定のキャッシュラインを置換するとき、メモリアクセスが最小になるように、識別されたバイト集合のみを、関連づけられるメモリへ書き込む。

【0011】本発明の前記および他の目的、特徴、利点は以下の詳細な説明から明らかとなる。

【0012】

【発明の実施の形態】ここで図1を参照すると、そこには本発明のライトバックキャッシュを含むデータ処理システム10が示される。データ処理システム10はプロセッサ12、オンボードのレベル1(L1)キャッシュ14、オプションのレベル2(L2)キャッシュ16、メモリ18、システムバス20、プロセッサバス22、およびメモリコントローラ24を含む。

【0013】プロセッサ12はソフトウェア命令を実行する。ソフトウェア命令はメモリ18に記憶されるか、システムバス20へ接続された他の装置からプロセッサ12によって受け取られる。メモリ18への反復アクセスによって遅延が生じるのを避けるため、メモリ18から検索された命令とデータは、それらが再び要求される可能性があるため、レベル1(L1)キャッシュ14またはレベル2(L2)キャッシュ16(もしあれば)に記憶される。プロセッサ12は、縮小命令セットコンピュータ(RISC)技術に従って動作する、IBM Microelectronicsから入手できるPowerPC系マイクロプロセッサの1つであることが望ましい。しかし、80x86、Intel Pentium、または他の適当なプロセッサであってもよい。

【0014】レベル1(L1)キャッシュ14は小型の(たとえば16Kバイト)の高速メモリを含む。このメモリに対しては、プロセッサ12内のレジスタおよび実行ユニットが、高速のプロセッサバス22およびシステムバス20を介してメモリ18から検索された最新のデータと命令へ迅速にアクセスすることができる。前述したように、データをメモリ18へ書き戻すために使用されるシステムバス20のバンドウィズを少なくすることによってデータ処理システム10のパフォーマンスを最大にするため、レベル1(L1)キャッシュ14はライトバックキャッシュとして構成するのが望ましい。しかし、レベル1(L1)キャッシュ14の構成は、そこに含まれるデータへのアクセスがキャッシュライン単位でライトスルーで行われるかライトバックで行われるかをソフトウェアで決めることができる。レベル1(L1)キャッシュ14はデータと命令の双方を記憶する一体型キャッシュとして示されているが、当業者に明らかのように、データキャッシュと命令キャッシュの2つに分けてもよい。

【0015】プロセッサ12は、ダッシュ線で示されるように、オプションのレベル2(L2)キャッシュ16に接続することができる。レベル2(L2)キャッシュ16は、レベル1(L1)キャッシュ14と同じく高速メモリを含み、メモリ18内に記憶されたデータと命令のサブセットを含む。データアクセスの効率を最大限にするため、レベル2(L2)キャッシュ16のサイズはレベル1(L1)キャッシュ14よりも大きく、またライトバックキャッシュとして構成するのが望ましい。レベル2(L2)キャッシュ16は、専用バス26を介してプロセッサ12へ接続されるように示されているが、当業者に明らかのように、プロセッサ12のオンボードとし、インラインまたはルックアサイド構成でプロセッサバス22に接続するか、システムバス20へ接続することができる。

【0016】メモリ18はランダムアクセスメモリ(RAM)から構成され、このランダムアクセスメモリは、図示されるようにメモリコントローラ24に関連づけられている。メモリコントローラ24はデータの記憶と検索を行うための読み出し能動信号と書き込み能動信号を生成し、またプロセッサ12によって使用される論理アドレスをメモリ18内の物理アドレスへ対応づけるアドレス変換機構を含んでいる。当業者に明らかであるように、メモリ18は多数の揮発性メモリモジュールから構成してよく、このメモリモジュールは、電力がデータ処理システム10へ供給されている間、データと、オペレーティング・システムおよびアプリケーションソフトウェアのセグメントとを記憶する。さらに、当業者に明らかであるように、システムバス20にはメモリ18のほか各種の装置を接続して、プロセッサ12、レベル1(L1)キャッシュ14、レベル2(L2)キャッシュ

16、およびメモリ18と交信を行わせることができる。しかし、このような装置は本実施例の理解には必要であり、説明を簡単にするために省いてある。

【0017】ここで図2を参照すると、そこには図1で示したデータ処理システムに使用されるライトバックキャッシュが示される。図2に示されたライトバックキャッシュはレベル1(L1)キャッシュ14として示されているが、この図はレベル2(L2)キャッシュ16へも同じく適用可能であることを理解されたい。図2において、レベル1(L1)キャッシュ14は多数のキャッシュライン30を含む2ウェイセット関連キャッシュである。キャッシュライン30の各々は、2つのウェイの各々で4つのバスバケットP0-P3を含んでいる。バスバケットP0-P3の各々は、1つまたは複数のデータバイトを含み、またプロセッサバス22およびシステムバス20の幅に等しいパケットサイズであることが望ましい。したがって、プロセッサバス22およびシステムバス20が64ビットのバスであれば、バスバケットP0-P3の各々は8つのデータバイトを含むことが望ましい。

【0018】通常のキャッシュと同じく、データは、データを記憶するメモリアドレスのインデックス部分によって、レベル1(L1)キャッシュ14の特定のキャッシュライン30へマップされる。特定のメモリアドレスに関連づけられたデータは、レベル1(L1)キャッシュ14に記憶されたとき、アドレスタグ32によって、同じインデックス部分をもつ他のメモリアドレスに関連づけられたデータから識別される。アドレスタグ32は、データが記憶されているメモリ18内の物理アドレスの高順位ビットから構成されることが望ましい。

【0019】さらに、レベル1(L1)キャッシュ14は整合プロトコルビット34を含む。これらのビットは、実施例ではMEI(Modified, Exclusive, Invalid)キャッシュ整合プロトコルで使用される。当業者に明らかであるように、このプロトコルの代わりに他のキャッシュ整合プロトコル、たとえばMESI(Modified, Exclusive, Shared, Invalid)プロトコルをレベル1(L1)キャッシュ14内で使用できる。整合プロトコルビット34は、1つの排他ビット(E)37、1つの無効ビット(I)36、および各キャッシュライン30の各ウェイに関連した4つの変更ビット38(M0-M3)を含む。キャッシュライン30の特定のウェイに関連した排他ビット37がセットされているとき、それはウェイ内に記憶されたデータブロックがレベル1(L1)キャッシュ14にのみ含まれ(レベル2(L2)キャッシュ16には含まれず)、メモリ18と整合していることを示す。キャッシュライン30の特定のウェイに関連した無効ビット36がセットされているとき、それは対応するアドレスタグ32に関連したデータ

がレベル1(L1)キャッシュ14内に存在していないことを示す。たとえば、データ処理システム10内で第2バスマスタによってメモリ18内の同じ場所にキャッシュ可能な読み取りアクセスまたは書き込みアクセスが行われたとき、スヌープヒット(snoop hit)が起こると無効ビット36がセットされる。変更ビット38(M0-M3)の各々は、対応するキャッシュラインウェイ内のバスバケットP0-P3の特定の1つに関連づけられている。変更ビットM0-M3は、対応するバスバケットP0-P3がメモリ18に関して変更されたときにセットされる。したがって、変更ビットM0-M3の1つがセットされているとき、それは対応するバスバケット内のデータがレベル1(L1)キャッシュ14内でのみ有効であることを示し、そのデータが最終的には整合性を維持するためにメモリ18へ書き戻されなければならないことを示す。

【0020】レベル1(L1)キャッシュ14はキャッシュ制御ロジック40を含む。キャッシュ制御ロジック40は、整合プロトコルビット34をセットまたはクリアすることによってキャッシュ整合プロトコルを実行するだけでなく、論理アドレスからアドレスタグを生成し、LRUアルゴリズムに従って置換のためのキャッシュラインを選択する。

【0021】ここで図3を参照すると、そこには図1のデータ処理システム10内でキャッシュの整合を維持する方法のフローチャートが示される。キャッシュアクセス要求がレベル1(L1)キャッシュ14によってプロセッサ12から受け取られたとき、プロセスがブロック50で始まる。プロセスはブロック50からブロック52へ進み、キャッシュ制御ロジック40が、オペランドデータの論理(実効)アドレスに基づいて、アドレスタグ(すなわち、オペランドデータの物理アドレスの高順位ビット)を生成する。さらに、キャッシュ制御ロジック40は、論理アドレスのインデックス部分を使用して、キャッシュアクセス要求を、指定されたキャッシュライン30へマップする。次に、プロセスはブロック52からブロック54へ進む。ブロック54は、キャッシュアクセス要求が読み取り要求か書き込み要求かを決定する。キャッシュアクセス要求が読み取り要求であると決定されると、プロセスはブロック54からブロック56へ進む。

【0022】ブロック56は、読み取り要求がヒットになるかミスになるかを決定する。ブロック56の決定は、ブロック52で生成されたアドレスタグを、指定されたキャッシュライン30の各ウェイに関連づけられたアドレスタグ32と比較することによって行われる。生成されたアドレスタグが、指定されたキャッシュライン30のウェイに関連づけられたアドレスタグ32の1つと一致すると、キャッシュヒットが起こり、プロセスはブロック57へ進む。ブロック57では、レベル1(L

1) キャッシュ14が、要求されたデータをプロセッサ12へ戻す。その後で、プロセスはブロック70で終了する。しかし、レベル1(L1)キャッシュ14で、読み取り要求がミスになると、プロセスはブロック64からブロック68までに進む。これらのブロックでは、以下説明するように、要求されたデータのメモリアドレスがマップされたキャッシュライン30のウエイの1つで置換が実行される。

【0023】ブロック54へ戻って、キャッシュアクセス要求が、指定されたキャッシュライン30への書き込みであることが決定されると、プロセスはブロック54からブロック58へ進む。ブロック58は、書き込み要求がヒットになるかミスになるかを決定する。ブロック58の決定は、ブロック52で生成されたアドレスタグを、指定されたキャッシュライン30の各ウエイに関連したアドレスタグ32と比較することによって行われる。生成されたアドレスタグが、アドレスタグ32の1つと一致すると、ヒットが起こり、プロセスはブロック60および62へ進む。ブロック60および62では、指定されたキャッシュライン30内の適切なバイトが更新され、更新されたバイトを含むバスケットに関連した変更ビット38がセットされる。したがって、ライトバックに必要なバスの使用を減らすために、レベル1(L1)キャッシュ14は、従来型のライトバックキャッシュよりも精細に、キャッシュ内のデータの変更を表示することができる。次にプロセスはブロック62からブロック70へ進み、そこでプロセスは終了する。

【0024】再びブロック58を参照すると、キャッシュアクセス要求がミスになると(すなわち、生成されたアドレスタグが、指定されたキャッシュライン30のウエイに関連したアドレスタグ32と一致しないという決定がなされると)、プロセスはブロック64からブロック68に進む。前述したように、ブロック64からブロック68は、指定されたキャッシュライン30の選択されたウエイのライトバックを示している。実施例では、置換されるウエイはLRUアルゴリズムに従って選択される。このLRUアルゴリズムは、単一のLRUビットをキャッシュライン30の各々に関連づけることによって実行することができる。しかし、当業者に明らかであるように、他の置換方式を使用することができる。

【0025】図3のブロック64では、キャッシュ制御ロジック40が、バスケットP0-P3のどれが変更されたかを識別するために、指定されたキャッシュライン30内で置換のために選択されたウエイに関連する変更ビット38を検査する。次に、プロセスはブロック66へ進む。ここでは、変更をマークされたバスケットのみをメモリ18へ書き込む。従来型のライトバックキャッシュとは対照的に、変更をマークされないバスケットはカーストアウトされるだけで、メモリ18へは書き戻されず、したがってライトバックに必要なバスの使

用は減少する。次に、プロセスはブロック68へ進む。ここでは、要求されたデータを含むデータブロックがレベル2(L2)キャッシュ16またはメモリ18からフェッチされ、そのデータブロックが、指定されたキャッシュライン30の選択されたウエイへ書き込まれる。さらに、バスケットP0-P3のいずれが、変更されたデータを含むかを表示するために、選択されたウエイに関連する変更ビット38がセットまたはクリアされる。したがって、選択されたウエイの置換が読み取りミスに陥る場合、選択されたウエイに関連する変更ビット38のすべてがクリアされる。しかし、書き込みミスによってウエイが置換される場合、バスケットP0-P3のいずれが変更されたデータを含むかを表示するために、変更ビット38の1つまたは複数がセットされる。その後で、プロセスはブロック70へ進み、そこで終了する。

【0026】これまでに説明したデータ処理システムの実施例から分かるように、変更された整合プロトコル情報をキャッシュラインの細分精度で保持することにより、ライトバックのバス使用が最小になる優れたライトバックキャッシュが提供される。整合プロトコル情報をバスケット単位で保持することは、スラッシングを最小にするために、低集合関連のキャッシュで特に利点がある。

【0027】実施例を詳細に説明したので、形式と細部については、本発明の精神と範囲を逸脱することなく各種の変更を加えてよいことが当業者に明らかであろう。

【0028】まとめとして、本発明の構成に関して以下の事項を開示する。

(1) 複数のバイト集合を含む複数のキャッシュラインであって、関連づけられるメモリ内のアドレスに関連づけられたデータを記憶するものと、上記キャッシュライン内の上記複数のバイト集合の中で、上記関連づけられるメモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むバイト集合を識別する整合標識と、上記キャッシュラインの特定のものを置換するとき、上記識別されたバイト集合のみを上記関連づけられるメモリへ書き込むキャッシュ制御ロジックと、を具備するライトバックキャッシュメモリ。

(2) 上記(1)において、上記整合標識が複数のビットを含み、該複数のビットの各々が上記複数のバイト集合の1つと関連づけられているライトバックキャッシュメモリ。

(3) 上記(1)において、上記バイト集合の各々が、ライトバックキャッシュメモリによって、上記関連づけられるメモリへ同時に書き込むことができる最大数のバイトを含むことを特徴とするライトバックキャッシュメモリ。

(4) 上記(1)において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有

し、該複数のウエイの各々が、その上記複数のバイト集合の中で変更されたバイト集合を識別する整合標識を有することを特徴とするライトバックキャッシュメモリ。

(5) 上記(1)において、プロセッサと関連づけられた上位レベルキャッシュが含まれていることを特徴とするライトバックキャッシュメモリ。

(6) 上記(1)において、上記キャッシュ制御ロジックが、LRUアルゴリズムに従って上記キャッシュラインの特定のものを置換することを特徴とするキャッシュメモリ。

(7) プロセッサと、上記プロセッサへ接続されたメモリと、上記メモリへ接続されたライトバックキャッシュメモリとを具備し、該ライトバックキャッシュメモリには、複数のバイト集合を含む複数のキャッシュラインであって、上記メモリ内のアドレスと関連づけられたデータを記憶するものと、上記キャッシュライン内の上記複数のバイト集合の中で、上記メモリ内の対応するアドレスに記憶されたデータとは異なったデータを含むバイト集合を識別する整合標識と、上記ライトバックキャッシュメモリ内の特定のキャッシュラインを置換するとき、上記識別されたバイト集合のみを上記メモリへ書き込むキャッシュ制御ロジックと、が含まれていることを特徴とするデータ処理システム。

(8) 上記(7)において、上記整合標識が複数のビットを含み、該複数のビットの各々が上記複数のバイト集合の1つに関連づけられていることを特徴とするデータ処理システム。

(9) 上記(7)において、上記バイト集合の各々が、上記ライトバックキャッシュメモリによって上記メモリへ同時に書き込むことができる最大数のバイトを含むことを特徴とするデータ処理システム。

(10) 上記(7)において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有し、該複数のウエイの各々が、そのバイト集合の中の変更されたものを識別する整合標識を有することを特徴とするデータ処理システム。

(11) 上記(7)において、上記ライトバックキャッシュメモリが上位レベルのキャッシュを含むことを特徴とするデータ処理システム。

(12) 上記(7)において、上記キャッシュ制御ロジックがLRUアルゴリズムに従って上記特定のキャッシュラインを置換することを特徴とするデータ処理システム。

(13) 複数のキャッシュラインを有するライトバックキャッシュメモリの中で整合性を維持する方法であって、関連づけられるメモリ内のアドレスに関連づけられたデータを、上記複数のキャッシュラインの中の特定のキャッシュラインに記憶するステップと、上記特定のキャッシュライン内の複数のバイト集合の中で、上記関連づけられるメモリ内の対応するアドレスに記憶されたデ

ータとは異なったデータを含むバイト集合を識別するステップと、上記ライトバックキャッシュメモリ内の上記特定のキャッシュラインを置換するとき、上記識別されたバイト集合のみを上記関連づけられるメモリへ書き込むステップとより成る、ライトバックキャッシュメモリ内で整合性を維持する方法。

(14) 上記(13)において、上記ライトバックキャッシュメモリがそれに関連づけられた複数のビットを有し、上記識別ステップでは、上記複数のビットの中で、識別されるバイト集合に関連づけられたビットがセットされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

(15) 上記(13)において、上記ライトバックキャッシュメモリと上記関連づけられるメモリがバスによって接続され、上記書き込みステップでは、識別されたバイト集合のためのバストランザクションのみが必要とされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

(16) 上記(13)において、上記複数のキャッシュラインの各々が、複数のバイト集合を含む複数のウエイを有し、該複数のウエイの各々が、その上記複数のバイト集合の中の変更されたものを識別する整合標識を有し、上記識別ステップでは、上記特定のキャッシュラインの選択されたウエイに含まれる上記複数のバイト集合の中の識別されるべきバイト集合に関連したビットがセットされることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

(17) 上記(13)において、上記特定のキャッシュラインがLRUアルゴリズムに従って置換されることを特徴とする、ライトバックキャッシュメモリ内で整合性を維持する方法。

【図面の簡単な説明】

【図1】少なくとも1つのライトバックキャッシュを含むデータ処理システムのブロック図である。

【図2】ライトバックキャッシュの実施例を示す図である。

【図3】ライトバックキャッシュ内でキャッシュの整合性を維持するために使用される方法を示すフローチャートである。

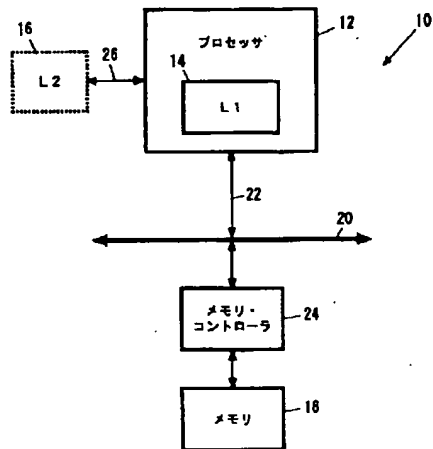
【符号の説明】

10	データ処理システム
12	プロセッサ
14	レベル1(L1)キャッシュ
16	レベル2(L2)キャッシュ
18	メモリ
20	システムバス
22	プロセッサバス
24	メモリコントローラ
26	専用バス
30	キャッシュライン

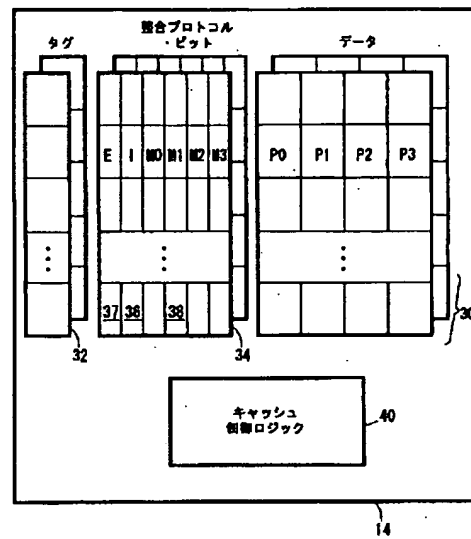
32 アドレスタグ
 34 整合プロトコルビット
 36 無効ビット
 37 排他ビット
 38 変更ビット
 40 キャッシュ制御ロジック
 50 ブロック
 52 ブロック
 54 ブロック

56 ブロック
 57 ブロック
 58 ブロック
 60 ブロック
 62 ブロック
 64 ブロック
 66 ブロック
 68 ブロック
 70 ブロック

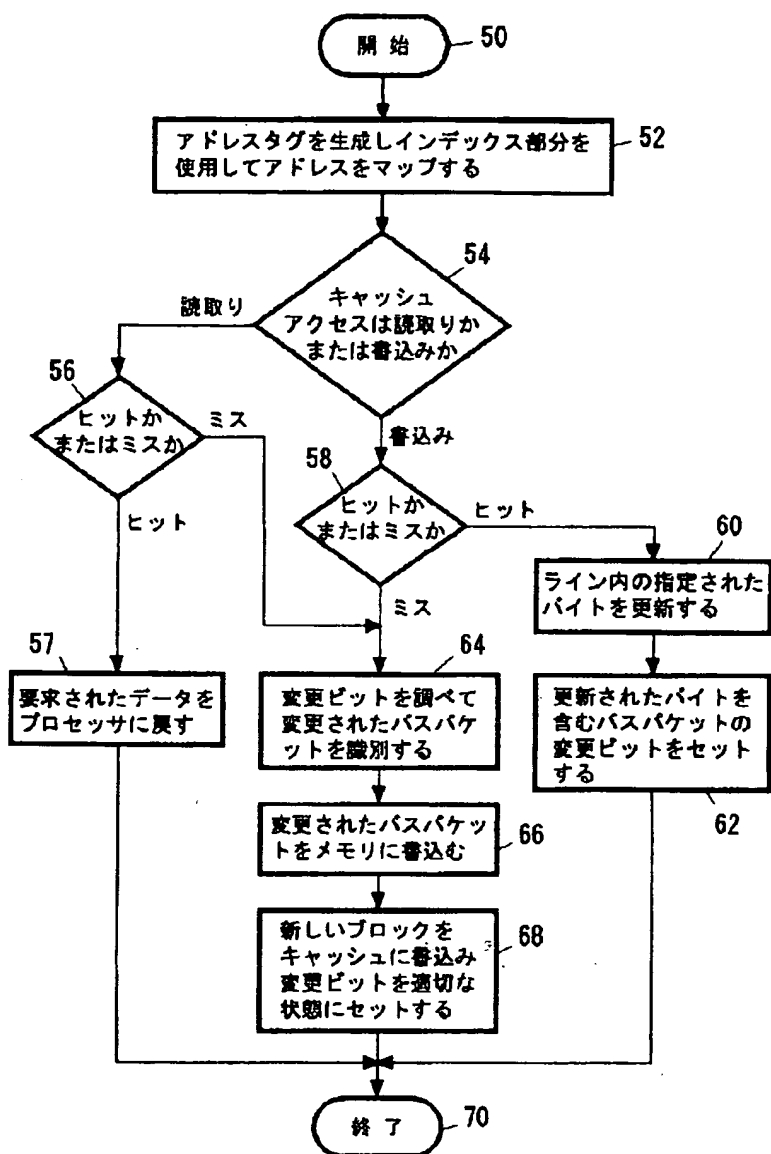
【図1】



【図2】



【図3】



フロントページの続き

(72)発明者 ソウミヤ・マリック
 アメリカ合衆国78729、 テキサス州オー
 スティン パートリッジ ベント ドライ
 ブ 13032